INFORMATION SCIENCES **73,** 273–284 (1993)

Addition as Fuzzy Mutual Entropy

BART KOSKO

Department of Electrical Engineering—Systems, Signal and Image Processing Institute, University of Southern California, Los Angeles, California 90089-2564

ABSTRACT

A sum of real numbers equals the mutual entropy of a fuzzy set and its complement set. A "fuzzy" or multivalued set is a point in a unit hypercube. Fuzzy mutual (Kullback) entropy arises from the logarithm of a unique measure of fuzziness. The proof uses the logistic map, a diffeomorphism that maps extended real space onto the fuzzy cube embedded in it. The logistic map equates the sum of a vector's components with the mutual entropy of two dual fuzzy sets. Diffeomap projection offers a new way to study the fuzzy structure of real algorithms.

1. THE THEOREM: ADDITION AS FUZZY MUTUAL ENTROPY

Any sum of real numbers $x_1, \ldots, x_n$ equals the fuzzy mutual entropy of fuzzy set F in the unit hypercube $[0,1]^n$:

$$\sum_{i=1}^{n} x_i = H(F/F^c) - H(F^c/F) \tag{1}$$

where F^c is the fuzzy set complement of F in the unit hypercube I^n. If you add two numbers, F and F^c lie in the unit square I^2. If you add three numbers, they lie in I^3, and so on up.

The proof of (1) maps the extended real space $\overline{R}^n = [-\infty, \infty]^n$ diffeomorphically onto the embedded unit hypercube I^n. The proof views the real numbers $x_1, \ldots, x_n$ as the components of the real vector $\mathbf{x}$ in $\overline{R}^n$ and maps $\mathbf{x}$ to a unique point or fuzzy set F in the fuzzy space I^n.

The mutual entropy terms $H(F/F^c)$ and $H(F^c/F)$ stem from the logarithm of the fuzziness of F. As discussed below this fuzziness depends on how much F resembles its complement fuzzy set F^c. In this sense we

655 Avenue of the Americas, New York, NY 10010 0020-0255/93/$6.00

can replace the two H terms in (1) with an entropy operator $\mathscr{H}$ applied to fuzzy set F:

$$\sum_{i=1}^{n} x_i = \mathscr{H}(F). \tag{2}$$

The operator $\mathscr{H}$ replaces each sum with the value of a map from fuzzy sets to real numbers.

The infinity "corners" of $-\infty$ and ∞ in $\bar{R}^n$ correspond to the $0-1$ vertices in I^n. The origin in $\bar{R}^n$ corresponds to the midpoint of I^n, the unique fuzzy set F such that $F = F^c = F \cap F^c = F \cup F^c$. The next three sections review the needed fuzzy information theory and develop the new measure of fuzzy mutual entropy. Section 5 proves (1).

2. FUZZY SETS AS POINTS IN HYPERCUBES: DEGREES OF SUBSETHOOD

Multivalence or "fuzziness" holds in sets and between sets. Fuzziness in a set defines *elementhood*, the degree a_i to which element x_i belongs to set A: $a_i = Degree(x_i \in A)$.

A standard or bivalent or "nonfuzzy" set A contains elements all or none. The membership degree a_i is 1 or 0, present or absent, in or out.

A multivalent set A contains elements to some degree. So a_i takes values in the unit interval $[0, 1]$. Black [1] called this multivalence "vagueness" and introduced vague sets or vague lists. Zadeh [12] called these vague or multivalued sets "fuzzy" sets and developed their algebra.

Fuzziness between sets defines *subsethood* [7–10], the degree $S(A, B)$ to which set A belongs to, or is a subset of, set B: $S(A, B) = \text{degree}(A \subset B)$. The sets A and B need not be fuzzy. If a fuzzy set A contains an element x_i to degree a_i, then $S(\{x_i\}, A) = a_i$. So subsethood subsumes elementhood. In the past the subsethood operator S has defined a bivalent operator in both fuzzy and nonfuzzy set theory: $S(A, B) = 0$ or 1. The multivalued subsethood operator can also assume the values $0 < S(A, B) < 1$.

The subsethood operator arises from the unique l^p-extension of the Pythagorean theorem [8] in n dimensions:

$$\|A - B\|^p = \|A - B^*\|^p + \|B^* - B\|^p \tag{3}$$

for $p \geqslant 1$, n-vectors A, B, and B^*, and with the norm

$$\|A\|^p = \sum_{i=1}^{n} |x_i|^p. \tag{4}$$

The usual Pythagorean theorem holds if $p = 2$. For any p there are 2^n vectors or sets B^* that satisfy (3). Then $b_i^* = a_i$ or b_i. For fuzzy sets these 2^n choices reflect the 2^n choices of picking any vertex of the unit hypercube I^n as the origin or empty set. Once picked, $b_i^* = \min(a_i, b_i)$. If A and B are bit vectors or regular nonfuzzy subsets of finite space $X = \{x_1, \ldots, x_n\}$, then this implies that B^* equals A intersect B, and the same holds for fuzzy subsets A and B of X:

$$B^* = A \cap B. \tag{5}$$

Suppose set or space X is finite with $X = \{x_1, \ldots, x_n\}$. Then the 2^n nonfuzzy subsets of X map to the 2^n bit vectors of length n. These map in turn to the 2^n corners of the unit hypercube I^n. This equates a set with a point in the Boolean n-lattice. We can also view fuzzy subsets of X as n-vectors with components in $[0, 1]$. Then each vector component a_i of fuzzy set $A = \{a_1, \ldots, a_n\}$ defines a fuzzy unit or *fit* [7] and A defines a fit vector. Fit value a_i measures the degree to which element x_i belongs to or fits in set A. This identifies A with a point on or in the unit hypercube I^n [8]. Fuzzy sets fill in the Boolean n-cube to give the solid hypercube I^n. The midpoint of the unit cube is the fit vector $F = (\frac{1}{2}, \ldots, \frac{1}{2})$ where each element x_i belongs to F as much as it belongs to its complement F^c. The usual set operations apply to fit vectors as Zadeh [12] proposed for fuzzy set functions: $A \cap B = (\min(a_1, b_1), \ldots, \min(a_n, b_n))$, $A \cup B = (\max(a_1, b_1), \ldots, \max(a_n, b_n))$, $A^c = (1 - a_1, \ldots, 1 - a_n)$. Suppose $A = (\frac{1}{3} \ \frac{3}{4})$ and $B = (\frac{1}{2} \ \frac{1}{3})$. Then

$$A \cap B = \left(\tfrac{1}{3} \quad \tfrac{1}{3}\right)$$

$$A \cup B = \left(\tfrac{1}{2} \quad \tfrac{3}{4}\right)$$

$$A^c = \left(\tfrac{2}{3} \quad \tfrac{1}{4}\right)$$

$$A \cap A^c = \left(\tfrac{1}{3} \quad \tfrac{1}{4}\right)$$

$$A \cup A^c = \left(\tfrac{2}{3} \quad \tfrac{3}{4}\right).$$

Note that $A \cap A^c \neq \emptyset$ and $A \cup A^c \neq X$ in this sample and for all fuzzy sets A. Aristotle's bivalent "law" of noncontradiction and excluded middle no longer hold. They hold only to some degree. They hold 100% only for the bit vectors at cube vertices. They hold 0% at the cube midpoint when $A = A^c$. For fit vectors between these extremes they hold only to some degree. The next section shows how the *overlap* term $A \cap A^c$ and *underlap* term $A \cup A^c$ give a unique measure of the fuzziness [5] or entropy of A.

If A and B are not fuzzy sets, then the 100% subsethood relation $A \subset B$ holds if and only if $a_i \leqslant b_i$ for all i. It still holds if A and B are fuzzy sets: $S(A, B) = 1$ iff $a_i \leqslant b_i$. Then all of B's 100% subsets define a hyperrectangle in I^n with a long diagonal that runs from the origin to point B. $S(A, B) = 1$ iff A lies in or on this hyperrectangle, the fuzzy power set of $B, F(2^B)$. $S(A, B) < 1$ iff A lies outside the hyperrectangle. The closer A lies to the hyperrectangle, the larger the value $S(A, B)$. The minimum distance lies between A and B^*, the 100% subset of B closest to A in *any* l^p metric [8]. This distance gives the l^p "orthogonal" projection of A onto $F(2^B)$ shown in Figure 1 and gives the term $\|A - B^*\|^p$ in the general l^p-Pythagorean theorem (3).

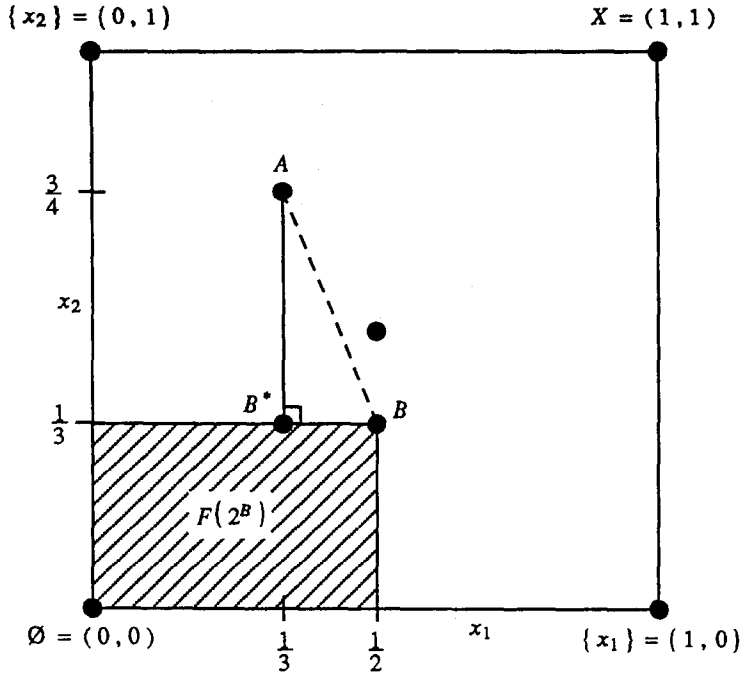

Fig. 1. Pythagorean geometry of the subsethood theorem of fuzzy sets.

The subsethood theorem follows from this orthogonal projection and unifies multivalued set theory. To see this first let $c(A)$ denote the *count* or cardinality of A:

$$c(A) = a_1 + a_2 + \ldots + a_n \tag{6}$$

$$= |a_1 - 0| + |a_2 - 0| + \ldots + |a_n - 0| \tag{7}$$

$$= l^1(\varnothing, A). \tag{8}$$

If $A = (\frac{1}{3}\ \frac{3}{4})$, then $c(A) = \frac{13}{12}$. The equalities (6)–(8) geometrize the count $c(A)$ as the l^1 or *fuzzy Hamming distance* between A and the origin or empty set $\varnothing$. The count extends the classical counting measure of combinatorics to fuzzy sets [6]: $c(A)$ equals the counting measure of A on nonfuzzy sets A. It gives the number of elements in A if A is finite and gives ∞ if A is infinite—if A maps one-to-one to one of its proper subsets. The subsethood measure $S(A, B)$ depends on the minimal distance $d(A, B^*)$. In the fuzzy Hamming metric this means $S(A, B) = 1 - l^1(A, B^*)/f(A)$. Boundary conditions on the empty set [8], [10] show that $f(A) = c(A)$. Since $B^* = A \cap B$, $l^1(A, B^*)$ equals the l^1 difference $[c(A) - c(\mathrm{A} \cap \mathrm{B})]$ shown in Figure 1. So $S(A, B) = 1 - [c(A) - c(A \cap B)]/c(A)$. This gives the *subsethood theorem*:

$$S(A, B) = \frac{c(A \cap B)}{c(A)}. \tag{9}$$

If $A = (\frac{1}{3}\ \frac{3}{4})$ and $B = (\frac{1}{2}\ \frac{1}{3})$, then $S(A, B) = \frac{2}{3} / \frac{13}{12} = \frac{8}{13}$ and $S(B, A) = \frac{2}{3} / \frac{5}{6} = \frac{4}{5}$. So B is more a subset of A than A is of B.

The derived ratio in (9) has the same form as the conditional probability $P(B/A)$. In general the event probability $P(A)$ is the degree to which the sample space X is a subset of its own subset or event A, $P(A) = S(X, A)$. This looks like the identity $P(A) = P(A/X)$. The subsethood theorem (9) also implies that the whole-in-the-part term $S(X, A)$ gives the relative frequency n_a/n if A denotes a bit vector with n_A 1s or successes and with $n - n_A$ 0s or failures: $S(X, A) = c(A \cap X)/c(X) = c(A)/c(X) = n_a/n$.

The subsethood theorem (9) also implies $S(\{x_i\}, A) = a_i$ since the singleton set $\{x_i\}$ maps to the unit bit vector $(0 \ldots 0\,1\,0 \ldots 0)$ with a 1 in the i^{th} slot and 0s elsewhere and since $A = (a_i, \ldots, a_n)$. Then $c(\{x_i\}) = 1$ and $c(\{x_i\} \cap A) = a_i$. So $S(\{x_i\}, A) = a_i$ and subsethood formally subsumes elementhood.

Maps between unit cubes define fuzzy systems; $S: I^n \to I^p$. Fuzzy systems associate output fuzzy sets with input fuzzy sets and so generalize if–then rules. Fuzzy systems are uniformly dense in the space of continuous functions [9]: a fuzzy system can approximate any real continuous (or Borel measurable) function on a compact set to any degree of accuracy. The fuzzy system contains fuzzy rules of the form IF $X = A$, THEN $Y = B$ that associate an output fuzzy set B with an input fuzzy set A. The rule defines a fuzzy Cartesian product $A \times B$ or patch in the input–output state space $X \times Y$. A fuzzy system approximates a function by covering its graph with patches and averaging patches that overlap. All the rules fire to some degree as in a neural associative memory [10]. The approximation theorem shows that finite discretizations of A and B suffice for the covering. So the patch or fuzzy Cartesian product $A \times B$ reduces to a fuzzy n-by-p matrix M or relation or point in I^{np}. Then M defines the system mapping M: $I^n \to I^p$ and the subsethood measure in (9) applies to M. In the same product space each fuzzy system is a subset to some degree of all other fuzzy systems. Then (11) below shows that each fuzzy system has a unique numerical measure of fuzziness [5] or entropy.

3. FUZZINESS AND ENTROPY

How fuzzy is a fuzzy set? A nonfuzzy set lies at a vertex of cube I^n and has 0% fuzziness. The cube midpoint P equals its own opposite, $P = P^c$, and it alone has 100% fuzziness. In between it varies. The fuzziness of set F grows as the distance falls between F and F^c—as F and F^c lie closer to the midpoint P.

This cube geometry motivates the ratio measure of fuzziness, $E(F) = a/b$, where a is the distance $l^1(F, F_{\text{near}})$ from F to the nearest vertex F_{near} and b is the distance $l^1(F, F_{\text{far}})$ from F to the farthest vertex F_{far}. A long diagonal connects F_{near} to F_{far}. The *fuzzy entropy theorem* [7] reduces this ratio to a ratio of counts:

$$E(F) = \frac{c(F \cap F^c)}{c(F \cup F^c)}. \tag{10}$$

If $F = (\frac{1}{3}\ \frac{3}{4})$, then $E(F) = \frac{7}{12} / \frac{17}{12} = \frac{7}{17}$. Figure 2 shows the fuzzy entropy theorem in the unit square.

The fuzzy entropy theorem (10) shows that the fuzziness of fuzzy set F depends on how much its overlap $F \cap F^c$ and underlap $F \cup F^c$ break Aristotle's laws of noncontradiction and excluded middle. Since the underlap $F \cup F^c$ always fully contains the overlap $F \cap F^c$, $S(F \cap F^c, F \cup F^c) = 1$,

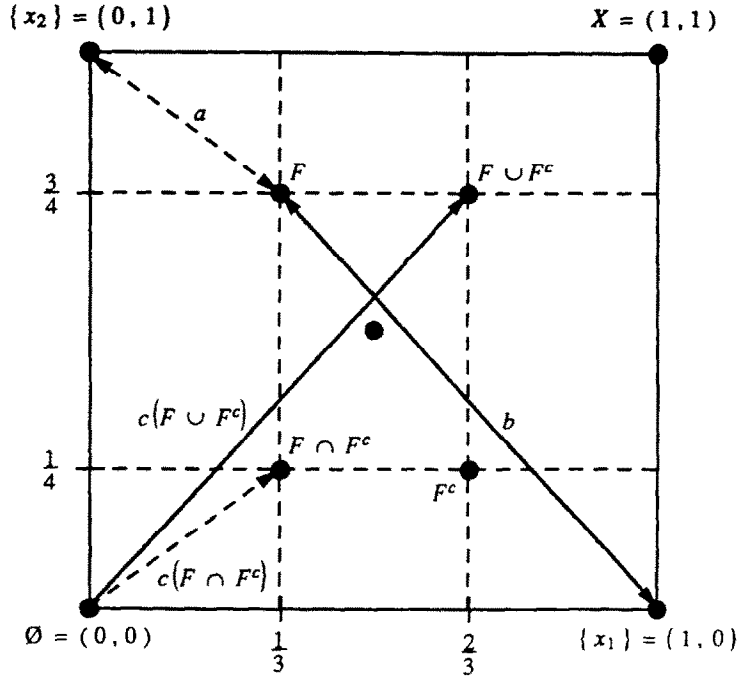

Fig. 2. Geometry of the fuzzy entropy theorem.

we might expect $E(F)$ to involve subsethood in the converse direction $S(F \cup F^c, F \cap F^c)$ when the part partially contains the whole. In fact (9) and (10) reduce fuzziness to subsethood:

$$E(F) = S(F \cup F^c, F \cap F^C). \tag{11}$$

The probabilistic entropy $H(P)$ [3, 5] holds for fit vectors on the simplex in I^n. Then

$$H(P) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i}, \tag{12}$$

and $c(P) = p_1 + \ldots + p_n = 1$. The fuzziness measure $E(P)$ differs from $H(P)$ for the same P. If no $p_j > \frac{1}{2}$, then $E(P) = 1/(n-1)$, and so $E(P)$ falls to zero as the cube dimension n grows to infinity. The uniform set $(\frac{1}{n}, \ldots, \frac{1}{n})$ belongs to this set of P vectors along with uncountably many

others. If some $p_j > \frac{1}{2}$, then $E(P) < 1/(n-1)$. So the uniform set maximizes $E(P)$ but does not uniquely maximizes it. So E differs from H.

Now consider how E resembles H. Consider the probability element p_i and the motivation for the logarithm measure (12) as the average information or entropy of a message or event: "information is inversely related to the probability of occurrence" [3]. The more improbable the event, the more informative the event if the event occurs. So information increases with $1/p_i$. The same intuition holds for monotone-increasing transforms of $1/p_i$. This includes the logarithmic transform $\log 1/p_i$ and only the logarithmic transform in the additive case. The weighted average over the system or alphabet gives the entropy as the expected information (12).

In the one-fit case $E(F)$ reduces to $f/(1-f)$ if $f \leqslant \frac{1}{2}$ and to $(1-f)/f$ if $f \geqslant \frac{1}{2}$. This ratio grows to 1 as f moves to the midpoint $\frac{1}{2}$ and falls to 0 as f moves to 0 or 1. The more vague or fuzzy the event, the more informative the event if it occurs. The operator E is subadditive on fuzzy sets since in a fuzzy space all events connect to one another to some degree. Integration also shows that $f/1-f$ and $1-f/f$ define a continuous probability density on $[0,1]$ if normalized by $\ln 4 - 1$. So far we have only reviewed fuzzy entropy. We now extend it to mutual entropy to set up the proof of the main theorem.

4. FUZZY MUTUAL ENTROPY

Fuzzy mutual entropy arises from a natural question: Why not take the logarithm of the unit fuzziness $f/(1-f)$? Any monotone transform will preserve its shape. So why not follow the probability example and use a logarithm? Then we can weight the log terms with the fit values and get a more proper measure of the entropy of a fuzzy set. The idea is to replace the intuition chain

$$p_i \to \frac{1}{p_i} \to \ln \frac{1}{p_i} \to \sum_i p_i \ln \frac{1}{p_i} \tag{13}$$

with the new fuzzy chain

$$f_i \to \frac{f_i}{1-f_i} \to \ln \frac{f_i}{1-f_i} \to \sum_i f_i \ln \frac{f_i}{1-f_i}. \tag{14}$$

The new fuzzy entropy term in (14) uses the natural logarithm to simplify the proof of the main theorem. The sum term defines a fuzzy mutual entropy.

For probability vectors P and Q in the I^n simplex, define the mutual entropy $H(P/Q)$ of P given Q [11] as

$$H(P/Q) = \sum_i p_i \ln \frac{p_i}{q_i}. \tag{15}$$

The mutual entropy measures distance in the simplex in the rough sense that $H(P/Q) = 0$ if $P = Q$, and $H(P/Q) > 0$ if $P \neq Q$. This follows from the Gibbs inequality [3]. Some stochastic learning automata and neural networks [4] minimize $H(P/Q)$ as the learning system's distribution P tries to estimate the distribution Q of the sampled environment. In the cube I^n, the fuzzy mutual entropy term in (14) is the usual mutual entropy $H(F/F^c)$ defined on fit vectors.

The sum of the fuzzy information units $\ln (f_i/1-f_i)$ splits into the mutual entropies of fuzzy sets F and F^c:

LEMMA:

$$\sum_i \ln \frac{f_i}{1-f_i} = H(F/F^c) - H(F^c/F). \tag{16}$$

Proof. Since $f_i + (1-f_i) = 1$,

$$\sum_i \ln \frac{f_i}{1-f_i} = \sum_i f_i \ln \frac{f_i}{1-f_i} + \sum_i (1-f_i) \ln \frac{f_i}{1-f_i} \tag{17}$$

$$= \sum_i f_i \ln \frac{f_i}{1-f_i} - \sum_i (1-f_i) \ln \frac{1-f_i}{f_i} \tag{18}$$

$$= H(F/F^c) - H(F^c/F). \quad \text{Q.E.D.} \tag{19}$$

The fuzziness measure in (10) shows that $E(F) = E(F^c)$. This reflects the 2^n-fold symmetry of the fuzzy cube I^n. But the mutual entropy operator is asymmetric. $H(F/F^c) = H(F^c/F)$ if $F = F^c$—if F and F^c lie

at the cube midpoint. The mutual entropy summands grow to infinity or zero as F and F^c move to cube vertices.

5. THE PROOF: DIFFEOMAPS BETWEEN REAL SPACES AND FUZZY CUBES

Fuzzy cubes map smoothly onto extended real spaces of the same dimension and vice versa. The 2^n infinite limits of extended real space $[-\infty,\infty]^n$ map to the 2^n binary corners of the fuzzy cube I^n. The real origin $\mathbf{0}$ maps to the cube midpoint. Each real point x maps to a unique fuzzy set F as Figure 3 shows.

A diffeomorphism f: $R^n \to I^n$ is a one-to-one and onto differentiable map f with a differentiable inverse f^{-1}. Different diffeomaps reveal different fuzzy structure of operations in real space. The theorem (1) follows from one of the simplest diffeomaps, the logistic map used in neural models [2, 10] to convert an unbounded real input x_i to a bounded signal or fit value f_i:

$$f_i = \frac{1}{1+e^{-x_i}}. \tag{20}$$

In extended real space $\overline{R}^n$ the logistic map applies to each term of vector $\mathbf{x}=(x_1,\ldots,x_n)$. Note that $f_i=0$ iff $x_i=-\infty$, $f_i=1$ iff $x_i=\infty$, and $f_i=\frac{1}{2}$ iff $x_i=0$. Each real $\mathbf{x}$ picks out unique dual fuzzy sets F and F^c in fuzzy space.

The proof of (1) follows from the lemma (16) and from the *inverse* of the logistic map (20):

$$x_i = f^{-1}(f_i) = \ln\frac{f_i}{1-f_i}. \tag{21}$$

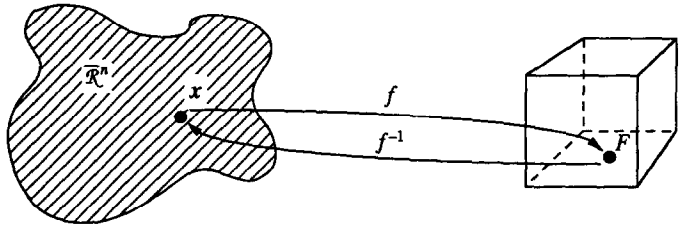

Fig. 3. Diffeomap from real space to fuzzy space.

So each real number is a unit of fuzzy information (14), the logarithm of the scalar measure of fuzziness (10). Sum over all n vector components x_i and apply the lemma (16) to prove (1) and (2):

$$\sum_i x_i = H(F/F^c) - H(F^c/F) \tag{22}$$

$$= \mathscr{H}(F) \tag{23}$$

in operator notation. Q.E.D.

The logistic map (20) also allows a direct proof for each term x_i:

$$x_i = x_i \frac{1+e^{-x_i}}{1+e^{-x_i}} \tag{24}$$

$$= \frac{1}{1+e^{-x_i}} \ln e^{x_i} - \frac{e^{-x_i}}{1+e^{-x_i}} \ln e^{-x_i} \tag{25}$$

$$= \frac{1}{1+e^{-x_i}} \ln \frac{1}{1+e^{-x_i}} \frac{1+e^{-x_i}}{e^{-x_i}}$$

$$- \frac{e^{-x_i}}{1+e^{-x_i}} \ln \frac{e^{-x_i}}{1+e^{-x_i}} (1+e^{-x_i}) \tag{26}$$

$$= f_i \ln \frac{f_i}{1-f_i} - (1-f_i) \ln \frac{1-f_i}{f_i} \tag{27}$$

since $f_i = 1/1+e^{-x_i}$ and $1-f_i = e^{-x_i}/1+e^{-x_i}$.

6. CONCLUSIONS

Addition or counting is the most basic operation in mathematics. It equals a basic operation in fuzzy space, the entropy map $\mathscr{H}$ that assigns a real number to each fuzzy set. This equality may seem odd since we have just begun to see the unit hypercube as a fuzzy space with its own set algebra and geometry. Diffeomap projection—or in some cases the weaker homeomorphic projection—can help show the fuzzy structure of real operations and algorithms. Future research may classify diffeomaps by how they preserve basic operations such as addition or how they carve the fuzzy cube into entropy regions or balls.

The fuzzy cube may also extend operations and algorithms in information theory. The fuzzy approximation theorem [9] converts continuous or

measurable systems into a finite number of fuzzy patches or points in large fuzzy cubes. The cube contains both the probability simplex that describes channel transmissions and the Boolean cube that describes all binary messages of a fixed length. An algorithm can dig through the cube from binary vertex to distant binary vertex rather than hop as a gray code from vertex to local vertex in cubes of high dimension. We can also view messages as balls of entropy or fuzziness in a cube. Ball diameter falls as the ball center moves from the vague midpoint to the clear corners. Diffeomaps can map real messages or systems into signal or noise balls that overlap in the fuzzy cube.

REFERENCES

1. M. Black, Vagueness: An exercise in logical analysis, *Philosophy of Science* 4:427–455 (1937).
2. S. Grossberg, *Studies of Mind and Brain*, Reidel, 1982.
3. R. W. Hamming, *Coding and Information Theory*, 2nd edition, Prentice Hall, 1986.
4. G. E. Hinton and T. J. Sejnowski, Learning and relearning in Boltzmann machines, in *Parallel Distributed Processing, Vol. I*, D. E. Rumelhart and J. L. McClelland, Eds., M.I.T. Press, 1986, pp. 282–317.
5. G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, 1988.
6. B. Kosko, Counting with fuzzy sets, *IEEE Trans. Pattern Anal. Machine Intell.* PAMI-8:556–557 (July 1986).
7. B. Kosko, Fuzzy entropy and conditioning, *Inform. Sci.* 40:165–174 (1986).
8. B. Kosko, Fuzziness vs. probability, *Int. J. General Syst.* 17(2):211–240 (1990).
9. B. Kosko, Fuzzy systems as universal approximators, in *Proc. IEEE Int. Conf. on Fuzzy Syst. 1992 (Fuzz'92)*, Mar. 1992, pp. 1153–1162.
10. B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice Hall, 1992.
11. S. Kullback, *Information Theory and Statistics*, Wiley, 1959.
12. L. A. Zadeh, Fuzzy sets, *Inform. Contr.* 8:338–353 (1965).

Received 5 September 1992; revised 19 November 1992